Generalized Bulgac-Kusnezov Methods for Sampling of the Gibbs-Boltzmann Measure

Leimkuhler, B10, 'Generalized Bulgac-Kusnezov Methods for Sampling of the Gobbs-Boltzmann Measure'
Physical Review E - Statistical, Nonlinear and Soft Matter Physics, vol 81, 026703

Generalized Bulgac-Kusnezov Methods for Sampling of the Gibbs-Boltzmann Measure

Benedict Leimkuhler

School of Mathematics and Maxwell Institute for Mathematical Sciences

University of Edinburgh

James Clerk Maxwell Building

Mayfield Road

Edinburgh EH9 2NX

United Kingdom

(Dated: January 12, 2010)

Abstract

A wide family of methods is described for sampling in the canonical ensemble. The Bulgac-Kusnezov method is generalized to include a more complicated coupling structure and stochastic perturbations. It shown that a controlled fluctuation of the potential surface or force field in a molecular model may be used as part of a sampling method (instead of the more standard friction/driving term). Numerical experiments demonstrate that the family includes novel methods that are effective for recovering canonical averages.

PACS numbers: 31.15.xv,82.60.-s

I. INTRODUCTION

The challenge of accurately sampling a complex molecular landscape is persistent and well-documented. Many interesting approaches have been suggested for addressing this problem (see, e.g. [1, 3] for some recent examples), yet it remains a major challenge for simulation based stategies in biomolecular modelling and computational materials science. There is strong demand for new concepts and methodology in this area. In particular it is valuable to have a comprehensive framework for modifying Hamiltonian dynamics in order that the canonical measure may be obtained from trajectory averages, i.e., thermostats, either based on dynamical or mixed stochastic-dynamical perturbation.

In this article, we describe a family of this type. Bulgac-Kusnezov [4] methods are closed form systems in extended phase space with (typically non-Hamiltonian) vector fields. This article provides a generalization of Bulgac-Kusnezov, with a more flexible coupling, and the addition of stochastic noise, that subsumes many existing thermostat methods, including methods such as Langevin dynamics and the recently proposed Nosé-Hoover-Langevin (NHL) method [6, 7] that introduce stochastic noise to stabilize the canonical measure.

As an example and to provide motivation for this new family, we are able to demonstrate some new types of methods, including a technique that allows thermal control based on activation of a force modification by an auxiliary variable. Effectively we replace the usual canonical sampling of a given system (by following perturbed trajectories in phase space) by a method based on trajectories in a potential energy surface is that modified dynamically (in such a way that canonical sampling is achieved). While we do not claim that the new approach is better than existing methods such as Langevin Dynamics, it is shown to be different and to have similar accuracy (with respect to averages) and convergence properties in a double well example.

II. CANONICAL MEASURE INVARIANT DYNAMICS

Our starting point is a Hamiltonian system

$$\dot{z} = J\nabla H(z) \tag{1}$$

on a $2N$ dimensional phase space M^{2N}, with J assumed to be a constant, nondegenerate, skew-symmetric matrix. The most obvious applications are to molecular dynamics in an

empirical potential U, with Hamiltonian (energy function) defined in terms of nuclear positions and momenta:

$$H = \frac{1}{2}\sum_{i=1}^{N} m_i^{-1} p_i^2 + U(q_1, q_2, \ldots, q_N),$$

but other applications are also of interest. Our goal in this article is to use dynamical paths to recover canonically weighted averages, i.e. spatial averages with respect to the Gibbs-Boltzmann (canonical) measure $d\mu_\beta = e^{-\beta H}\mathrm{d}^{2N}z$, where $\beta^{-1} = k_B T$ is the desired scaled temperature. (Our treatment could be extended to arbitrary smooth measures, following [8].) Note that the dynamical system (1) preserves the Hamiltonian, hence any function of the Hamiltonian, hence the canonical measure. The problem is that it provides no specificity, thus an initial distribution in phase space will typically decay to a collection of isolated components and averages taken along trajectories will not converge to canonical phase space averages.

In general, the idea of a dynamical thermostat is to replace the Hamiltonian dynamics (1) by a new system, usually in an extended phase space, designed so that averages with respect to the augmented system are easily mapped to canonical averages.

The method of Bulgac and Kusnezov (BK) [4] is of this type. Their proposal replaces the microcanonical system by

$$\dot{z} = J\nabla H(z) - \sum_{i=1}^{k} g_i'(\xi_i) F_i(z), \tag{2}$$

where, for $i = 1, 2, \ldots, k$,

$$\dot{\xi}_i = (\nabla_z H \cdot F_i - \beta^{-1}\nabla_z \cdot F_i)/\alpha_i, \tag{3}$$

with the functions $g_i : \mathbf{R} \to \mathbf{R}$ and $F_i : \mathbf{R}^{2N} \to \mathbf{R}^{2N}$ arbitrary smooth functions, and α_i, $i = 1, \ldots, k$ constant coefficients. It is easily demonstrated that under certain conditions on the expression in (3) (see [5]) this dynamics preserves an augmented canonical measure of the form

$$\mathrm{d}\tilde{\mu}(z, \xi) = \rho_\beta \times e^{-\beta\sum_{i=1}^{k} g_i(\xi_i)}\mathrm{d}^{2N}z\mathrm{d}^k\xi \equiv \tilde{\rho}\mathrm{d}^{2N}z\mathrm{d}^k\xi. \tag{4}$$

Specifically, one constructs the Liouville operator of the extended system

$$\mathcal{L}\rho = -\nabla_z \cdot (\rho\dot{z}) - \nabla_\xi \cdot (\rho\dot{\xi}), \tag{5}$$

and it is then a straightforward calculation to show that $\mathcal{L}\tilde{\rho} = 0$. This means that $\tilde{\rho}$ is stationary for the extended system. One assumes the functions g_i tend to $+\infty$

sufficiently rapidly, so that "integrating out" with respect to the auxiliary variables, we would hope to recover canonically weighted averages of functions of the physical variables. Unlike Hamiltonian dynamics (which also preserves the canonical measure), the BK method does not preserve arbitrary functions of the Hamiltonian, thus it achieves some additional specificity of the measure. However, there are no proofs of ergodicity; it is not known in which cases ρ^{aug} is the unique invariant density for the dynamical system and, in some specific cases, ergodicity is known to fail, see [12]. On the other hand, in practice, the method appears to reproduce canonical averages with some accuracy when the underlying physical dynamics is itself ergodic (or nearly so). It is thus common to work with dynamical thermostats under an assumption of ergodicity (even though this is unlikely to hold in the strict sense) in order to obtain approximation results useful in practical calculation.

The BK method includes some popular schemes for molecular dynamics such as Nosé-Hoover dynamics[9, 10]. For a system described by a Hamiltonian $H = H(q,p) = p^T M^{-1} p/2 + U(q)$, M a $N \times N$ constant mass matrix, the Nosé-Hoover method consists of the extension

$$\dot{q} = M^{-1}p, \tag{6}$$

$$\dot{p} = -\frac{\partial H}{\partial q} - \xi p, \tag{7}$$

$$\dot{\xi} = \alpha^{-1}\left[p^T M^{-1} p - N\beta^{-1}\right], \tag{8}$$

based on a single additional variable. Within the BK framework one can find some schemes which have the potential to accelerate sampling, including configurational thermostats (see e.g. [6, 11]).

Observe that, regardless of what choice is made for the functions g_i in (2)-(3), the average of g_i' is zero:

$$\langle g_i' \rangle = \frac{\int_{\mathbf{R}} g_i'(\xi_i) e^{-\beta g_i(\xi_i)} \mathrm{d}\xi_i}{\int_{\mathbf{R}} e^{-\beta g_i(\xi_i)} \mathrm{d}\xi_i} = 0, \tag{9}$$

upon integrating the numerator and using the fact that g_i tends to $+\infty$ as $\xi \to \pm\infty$.

III. GENERALIZED BULGAC-KUSNEZOV METHODS

We consider here extensions on the space $M^{2N} \times L^k$, defined by equations of the form

$$\dot{z} = u(z, \xi), \tag{10}$$

$$\dot{\xi} = v(z, \xi). \tag{11}$$

Assuming that (10)-(11) has an invariant measure $\mathrm{d}\tilde{\mu}$, we ask that

$$\int_{M^{2N} \times L^k} \varphi(T(z, \xi)) \mathrm{d}\tilde{\mu} = \int_{M^{2N}} \varphi(z) \mathrm{d}\mu_\beta, \tag{12}$$

for some appropriate choice of the transformation T. The family (10)-(11) includes Nosé-Hoover Chains [15], and other previous generalizations of Nosé-Hoover [14]. To see that the new formulation is more general than Bulgac-Kusnezov, we observe that it includes methods with complicated multi-variate nonlinear dependencies among the auxiliary variables, depending on the form of v, whereas Bulgac-Kusnezov methods have a simpler structure, $v = v(z)$. It is precisely this new freedom that we wish to exploit. In fact, the family (10)-(11) includes even the Hamiltonian Nosé-Poincaré scheme [13] which does not rely on a smooth product extension of the canonical measure; in this approach,

$$\tilde{H}^{\mathrm{NP}} = \xi_1 \left[H(q, p/\xi_1) + \frac{\xi_2^2}{2\alpha} + N\beta^{-1} \ln \xi_1 - \tilde{E} \right]. \tag{13}$$

and we define a Dirac measure in the extended energy:

$$\mathrm{d}\tilde{\mu} = \delta \tilde{H}^{\mathrm{NP}} \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}^N q \mathrm{d}^N p. \tag{14}$$

Then one easily shows that [13], with $L^k = \mathbf{R}_+ \times \mathbf{R}$, that

$$\int_{\mathbf{R}^{2N} \times L^k} f(q, p/\xi_1) \mathrm{d}\tilde{\mu} = \int_{\mathbf{R}^{2N}} f(q, p) \mathrm{d}\mu^\beta. \tag{15}$$

Under the essential ergodicity assumption, sampling of the extended H^{NP} Hamiltonian can be performed by viewing ξ_1 and ξ_2 as conjugate symplectic (canonical) variables in (13) and using Hamiltonian dynamics to generate "pseudo-microcanonical" sampling trajectories. Other, more complicated extended systems based on extended Hamiltonians have been constructed (see e.g. [16]) with the aim of improving the convergence of averages.

However, we are for the most part interested in smooth measures, i.e., we assume that, for some $g : \mathbf{R}^k \to \mathbf{R}$, our extended measure has the density $\tilde{\rho} = \rho_\beta \exp(-\beta g(\xi))$. From

direct computation, we must have

$$\mathcal{L}\tilde{\rho} = -\nabla_z \cdot (\tilde{\rho} u) - \nabla_\xi \cdot (\tilde{\rho} v) = 0. \tag{16}$$

Then it follows that

$$u \cdot \nabla_z H - \beta^{-1} \nabla_z \cdot u + \nabla_\xi g \cdot v - \beta^{-1} \nabla_\xi \cdot v \equiv 0. \tag{17}$$

As previously discussed, if g suitably chosen, it will be possible to compute averages with respect to the invariant density, and, in particular, we may obtain canonically weighted phase space averages by integrating out with respect to all the ξ_i.

A. Stochastic Perturbation

In order to provide ergodicity, we introduce diffusion in the density propagator by adding stochastic noise and dissipation to the dynamics of the extended system (10)-(11). We may allow noise to be introduced in either the physical or auxiliary variables:

$$\dot{z} = u(z, \xi) - \Gamma_z z + \sqrt{2\beta^{-1}} \Gamma_z^{1/2} \dot{W}_z, \tag{18}$$

$$\dot{\xi} = v(z, \xi) - \Gamma_\xi \xi + \sqrt{2\beta^{-1}} \Gamma_\xi^{1/2} \dot{W}_\xi, \tag{19}$$

where W_z and W_ξ are $2N$ and k-dimensional vectors of independent Wiener processes and Γ_z, Γ_ξ are matrices which define the coupling of the stochastic terms to the physical and auxiliary variables. Typically the coupling matrices would be taken to be constant diagonal matrices, or projections onto the physical momenta, as in the case of Langevin dynamics, or even a scalar; there are many alternatives. Γ_z might be obtained from physical principles to represent friction needed to represent neglected degrees of freedom, on the other hand Γ_ξ is likely to consist of artificial parameters.

If the condition (17) holds, and if the stochastic part

$$(-\Gamma_z z + \sqrt{2\beta^{-1}} \Gamma_z^{1/2} \dot{W}_z, -\Gamma_\xi \xi + \sqrt{2\beta^{-1}} \Gamma_\xi^{1/2} \dot{W}_\xi) \tag{20}$$

is chosen to preserve the reduced canonical measure with respect to each set of variables, then (18)-(19) preserves the canonical measure. This follows directly from the linear structure of the Fokker-Planck equation.

B. Nosé-Hoover-Langevin Thermostat

As an example of an existing scheme that combines dynamical extension and stochastic perturbation in the auxiliary variables, we mention the Nosé-Hoover-Langevin (NHL) method studied in [6, 7]. We assume we have a molecular system for which the energy may be written $H(q,p) = p^T M^{-1} p/2 + U(q)$. A variant of the Nosé-Hoover-Langevin method may be written

$$\dot{q} = M^{-1}p \tag{21}$$

$$\dot{p} = -\nabla \bar{U} - \epsilon\xi p \tag{22}$$

$$\dot{\xi} = p^T M^{-1} p - N\beta^{-1} - \gamma\xi + \sqrt{2\beta^{-1}\gamma\epsilon^{-1}}\dot{W}, \tag{23}$$

where W is a (scalar) Wiener process. This method includes a Nosé-Hoover-like control law to control kinetic energy; it replaces the artificial use of a thermostat chain [15] by a single stochastic process. The augmented distribution has the form $\tilde{\rho} = \rho_\beta e^{-\beta\epsilon\xi^2/2}$.

IV. SAMPLING METHODS BASED ON FORCE MODIFICATION

It would be desirable to be able to introduce a more flexible modification of the force field in order to provide flexibility in the way equilibrium is reached in molecular simulation. Such a mechanism could be valuable in the design of enhanced sampling strategies. Here we present a preliminary outline of such a technique which is based on viewing the given force field as a slice of a projected force field in an extended configurational space. The extra freedom may provide routes to avoid ergodicity barriers.

Consider replacing the conservative force $F(q) = -\nabla U(q)$ in a molecular system by a function $\tilde{F}(q,\xi)$, where ξ is driven by an (artificial) dynamical process. The equations of motion would be

$$\dot{q} = M^{-1}p, \tag{24}$$

$$\dot{p} = \tilde{F}(q,\xi), \tag{25}$$

$$\dot{\xi} = h(q,p,\xi). \tag{26}$$

Assuming a simple extension of the canonical density, e.g. $\tilde{\rho} = \rho_\beta e^{-\beta\xi^2/2\mu}$, we may easily derive the following solution for h:

$$h(q,p,\xi) = \beta e^{\beta\xi^2/2\mu} \int_0^\xi e^{-\beta s^2/2\mu} \Delta\tilde{F}(q,s) \cdot (M^{-1}p) ds, \tag{27}$$

where $\Delta\tilde{F} = \tilde{F} - F$. (This is obtained by writing out the differential equation satisfied by h, then solving it in the standard way using an integrating factor.)

One possibility is to imagine the graph of $U(q)$ as being embedded within a smoothed landscape in the extended variable. For example, we could set $\tilde{U}(q,\xi) = \sigma(\xi)U(q)$, then take $\tilde{F} = -\nabla_q \tilde{U}$, in which case Eq. (27) reduces to

$$h(q,p,\xi) = -\beta\omega(\xi)\nabla_q U(q)\cdot(M^{-1}p), \tag{28}$$

for a scalar function ω. The relationship between ω and σ being

$$\sigma = \omega' - \beta\mu\xi\omega + 1. \tag{29}$$

For some choices of σ we may invert this relationship to derive the appropriate choice of ω. For example take the function

$$\sigma(\xi) = 1 - C_2 \arctan^2(C_1\xi^2), \tag{30}$$

for suitable constants C_1, C_2. In this case it is possible to embed the physical potential in landscape that could potentially enable more alternative routes for exploration (see Figure 1). Here ω is recovered by inverting (29):

$$\omega(\xi) = e^{\mu\beta\xi^2/2\mu}\int_0^\xi e^{-\beta s^2/2\mu}\sigma(s)ds. \tag{31}$$

For polynomial σ a recurrence is available to solve this integral exactly, but in general cases this would need to be done by numerical quadratures. Experiments with this scheme showed numerical instability due to the presence of the rapidly growing exponential term.

The simplest nontrivial analytical solution is

$$\sigma(\xi) = \beta\mu\xi^2, \qquad \omega(\xi) = -\xi. \tag{32}$$

Ergodicity in the EFM method may be a practical issue, since EFM only receives contact with stochastic perturbation via the dependence of the force field on ξ. If $\tilde{F}_\xi$ vanishes in a certain region, the system must rely on mixing present in the physical dynamics. The system can be combined with additional thermostatting devices. For example, (24)-(26) would likely be adapted to include a stochastic term to drive the auxiliary variable to equilibrium:

$$\dot{q} = M^{-1}p, \tag{33}$$

$$\dot{p} = \tilde{F}(q,\xi), \tag{34}$$

$$\dot{\xi} = h(q,p,\xi) - \gamma\xi + \sqrt{2\gamma\beta^{-1}}\dot{W}. \tag{35}$$

It may also be desirable to complement the given scheme by Langevin dynamics in the physical variables or a Nosé-Hoover-Langevin thermostat. We refer to such a combination as the Embedded Force Method.

V. NUMERICAL EXPERIMENTS

We present in this section simple numerical experiments to examine some instances of the generalized Bulgac-Kusnezov schemes against the "gold standard" of stochastic sampling methods, Langevin dynamics. The goal is to compare the trajectories obtained from the different methods in order to gain insight into the approach to equilibrium.

Comparisons were made between the Embedded Force Method (of the previous section), Nosé-Hoover-Langevin, and Langevin Dynamics. In running comparisons, we have had to select a number of parameters (stepsizes, coupling coefficients) in each method. It is challenging to study all values of the parameters simultaneously for all methods. As our goal is just to test the recovery of canonical averages in a more flexible framework, such an exercise is in any event not particularly enlightening (future work will attempt to establish the usefulness of the described methodology in applications such as biomolecular modelling). The limited results presented here therefore cannot be taken to provide a comprehensive comparison of the different methods, but they do provide "proof of concept" for the novel approach presented in this article. In our experiments, we studied a double well potential,

$$U(q) = (q^2 - 1)^2. \tag{36}$$

We set $kT = 0.15$; as the barrier is of height 1, this provided a reasonably challenging example for studying the different methods. The system was initialized using random points in the left basin and with a small initial velocity. The challenge was then to observe accurate sampling of both basins.

All the formulations under study involve stochastic perturbations. Several recent articles have discussed the numerical treatment of stochastic dynamics used in molecular dynamics [18, 19], but consensus on the best method currently available has not been reached. In comparisons, we used the Stochastic Position Verlet (SPV) method from [19]. A variety (three) other recently proposed methods for Langevin dynamics were implemented and tested but in the experiments SPV proved be the best of the methods

and was certainly adequate for our purposes. The SPV method is given in the appendix.

For the Embedded Force Method of Section III, we chose

$$\tilde{U}(q,\xi) = \begin{cases} U(q), & |q| \geq 1 \\ (1-\beta\xi^2)U(q), & |q| < 1 \end{cases} \tag{37}$$

This leads to

$$h(q,p,\xi) = \begin{cases} 0, & |q| \geq 1 \\ \beta\xi U'(q)p, & |q| < 1 \end{cases}, \tag{38}$$

following the method of Section III. In the Embedded Force Method, we augmented the equations by a stochastic perturbation of ξ, as in (33)-(35) and we also added a Nosé-Hoover-Langevin thermostat to enhance ergodicity (as it possible to otherwise to find some orbits for which the thermostatting variables do not interact at all with the physical ones) however, we chose a very small value of ϵ, $\epsilon = 0.001$ in order to verify that the performance was not just a consequence of the NHL device. (We verified that with $\epsilon = 0.001$ in the pure NHL method, there are almost no crossings between basins.) The random friction term was fixed for all noise processes in all runs at $\gamma = 1.0$. The stepsize for all simulations was fixed at $h = 0.05$.

The first observation is that the three methods produce trajectories with very different characteristics. Although all three methods incorporate stochastic perturbation, the Langevin trajectories are much less smooth than those produced by the other two methods. The NHL trajectories are smooth and the EFM dynamics behaves similarly to Langevin dynamics when $p \approx 0$. Representative trajectories are shown in Figure 2. We can also see from these figures that NHL and EFM appear to provide a poorer sampling of the saddle point compared to Langevin Dynamics.

All three methods produce correct sampling of the canonical measure. For each method we computed 7 runs of 50M timesteps, reducing sampling errors to less than 5% for each method. Configurational sampling (comparing histograms of positions against the predicted densities) in one sampling run is shown for each method in Figure 3.

Finally, we also looked at convergence of the methods in terms of the occupancy time of the right basin. Computing 96 trajectories of length 5M for each method we calculated means and standard deviations. Figure 4 shows the graphs of the mean (bold solid) and standard deviation (bold dashed) curves against light curves showing the behavior for each individual run. Although differences among the methods are somewhat subtle, it seems that EFM eliminated outliers present in both the other methods and

had, therefore, a slightly smaller standard deviation. The reader is again cautioned that different parameter choices and different problem choices may impact the relative performance of the methods.

VI. CONCLUSION

We have demonstrated that the Bulgac-Kusnezov framework for sampling the canonical ensemble may be generalized to include a more complicated interaction among artificial variables and/or stochastic perturbations. We have also obtained, within this framework, a new method that thermally activates part of the force field in order to achieve canonical sampling. The method has been tested and compared with Langevin dynamics, to a limited extent, and performs well in sampling of an uneven double well, providing proof of concept. The obvious challenge is to employ this new method to thermostat some more relevant systems arising in materials or biological applications.

Acknowledgements. The author appreciates the support of the National Science Foundation funded Institute for Mathematics and Its Applications (IMA) in Minneapolis and the Warwick Mathematics Institute, both of which accommodated him during the preparation of this paper. The author further acknowledges the encouragement and helpful comments of Andrew Stuart and Mike Allen (both of Warwick) and the referees.

Appendix A

For Langevin dynamics, we obtained good results with the Stochastic Position Verlet (SPV) method [19], which is a splitting method which reduces to Verlet in the absence of stochastic noise ($\gamma = 0$). This scheme is shown to give second order accuracy for moments. The formulas for the SPV step are as follows:

$$\begin{aligned} q &= q + \frac{1}{2} h M^{-1} p \\ p &= \exp(-h\gamma) p - \frac{1 - \exp(-h\gamma)}{\gamma} \nabla U + \sqrt{\theta(1 - \exp(-2\gamma h))} M^{1/2} R \\ q &= q + \frac{1}{2} h M^{-1} p \end{aligned}$$

This method is second order for moments. Here $\theta = kT$.

The above methods are *quasisymplectic* in the sense of [20].

For the Embedded Force Method (EFM) and Nosé-Hoover-Langevin, the timestep was defined by

$$
\begin{aligned}
p &= p - \frac{1}{2}h\nabla(\bar{U}(q) + \alpha\hat{U}(q)) - \frac{1}{2}h(1-\alpha)\theta^{-1}\xi^2\nabla\hat{U}(q) \\
q &= q + \frac{1}{2}hM^{-1}p \\
p &= e^{-h\epsilon\eta/2}p \\
\xi &= \exp(-\frac{1}{2}h\theta^{-1}p\cdot M^{-1}\nabla\hat{U}(q))\xi \\
\eta &= \eta + \frac{h}{2}(p^TM^{-1}p - N\theta) \\
\xi &= \exp(-h\gamma)\xi + \sqrt{\theta\mu^{-1}(1-e^{-2\gamma h})}R_\xi \\
\eta &= \exp(-h\gamma)\eta + \sqrt{\theta\mu^{-1}(1-e^{-2\gamma h})}R_\eta \\
\eta &= \eta + \frac{h}{2}(p^TM^{-1}p - N\theta) \\
\xi &= \exp(-\frac{1}{2}h\theta^{-1}p\cdot M^{-1}\nabla\hat{U}(q))\xi \\
p &= e^{-h\epsilon\eta/2}p \\
q &= q + \frac{1}{2}hM^{-1}p \\
p &= p - \frac{1}{2}h\nabla(\bar{U}(q) + \alpha\hat{U}(q)) - \frac{1}{2}h\theta^{-1}\xi^2\nabla\hat{U}(q)
\end{aligned}
$$

where R_ξ and R_η are two independent standard normally distributed random numbers, γ_ξ and γ_η) are constants controlling the two noise processes, and $\sigma_\xi = \sqrt{2\beta^{-1}\gamma_\xi}$, $\sigma_\eta = \sqrt{2\beta^{-1}\gamma_\eta}$. The parameters ϵ and α control which method this is: if $\alpha = 1$ we have Nosé-Hoover-Langevin. With $\epsilon = 0$, we have a pure EFM thermostat. For $\epsilon \neq 0$, $\alpha \neq 1$ we have a combination of the two methods.

---

[1] Z. Zhu, M. E. Tuckerman, S. O. Samuelson and G. J. Martyna, *Phys. Rev. Lett.* **88**, 100201 (2002).

[2] L. Rosso, P. Minary, Z. Zhu and M. E. Tuckerman, *J. Chem. Phys.* **116**, 4389 (2002).

[3] A. Laio and F.L.Gervasio, *Rep. Prog. Phys.* **71**, 126601 (2008)

[4] A. Bulgac, and D. Kusnezov, *Phys. Rev. A* **42**, 5045 - 5048 (1990)

[5] O. Jepps, G. Ayton, D. Evans, *Phys. Rev. E* **62**, 4757-4763 (2000)

[6] A. Samoletov, C. Dettmann, and M. Chaplain, *J. Stat. Phys.* **128**, 1321-1336 (2007)

[7] B. Leimkuhler, E. Noorizadeh, and F. Theil, *J. Stat. Phys.* **135**, 261-277 (2009)

[8] E. Barth, B. Laird, and B. Leimkuhler, *J. Chem. Phys.* **118**, 5759-5767 (2003)

[9] S. Nosé, *J. Chem. Phys.* **81**, 511-519 (1984)

[10] Hoover, W., *Phys. Rev. A.* **31**, 1695-1697 (1985)

[11] K. Travis, and C. Braga, *J. Chem. Phys.* **128**, 014111 (2008)

[12] F. Legoll, M. Luskin and R. Moeckel, *Arch. Rat. Mech. Anal.* **184**, 449-463 (2007)

[13] S. Bond, B. Leimkuhler, and B. Laird, *J. Comput. Phys.* **151**, 114-134 (1999)

[14] A.C. Branka and K.W. Wojciechowski, *Phys. Rev. E* **62**, 3281 - 3292 (2000)

[15] G.J. Martyna, M.L. Klein, and M. Tuckerman, *J. Chem. Phys.* **97**, 2635-2643 (1992)

[16] B. Leimkuhler and C. Sweet, *SIAM J. Appl. Dyn. Syst.* **4**, 187-216 (2005)

[17] J.C. Mattingly and A.M. Stuart, *Markov Process. Related Fields*, **8**, 199–214 (2002)

[18] E. Vanden-Eijnden and G. Ciccotti, *Chem. Phys. Lett.* **429**, 310-316 (2006)

[19] S. Melchionna, *J. Chem. Phys.* **127**, 044108 (2007)

[20] G.N. Milstein and N.V. Tretyakov, *IMA J. Numer. Anal.***23**, 593-626 (2003)

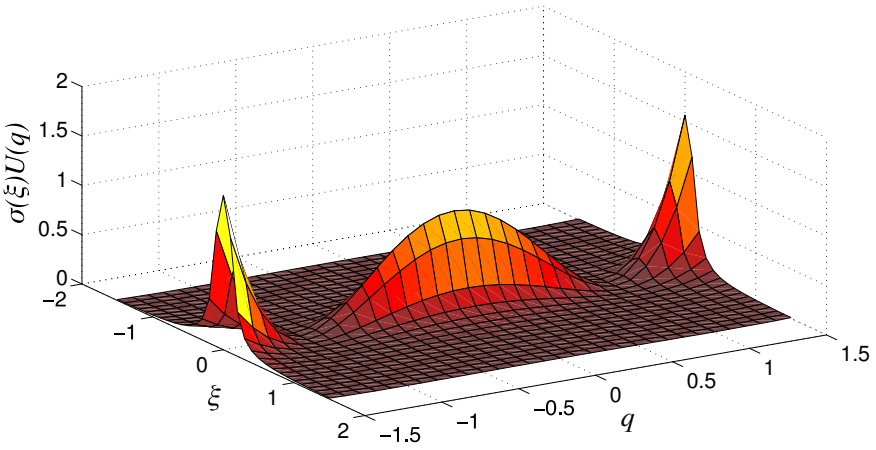

FIG. 1: Graph of $\tilde{U}(q,\xi) = \sigma(\xi)U(q)$ where U is a double well and $\sigma(\xi) = [\frac{2}{\pi}\arctan(100\xi^2)]^2$. The slice along $\xi = 0$ represents the original potential, which is also a shift of the effective free energy.

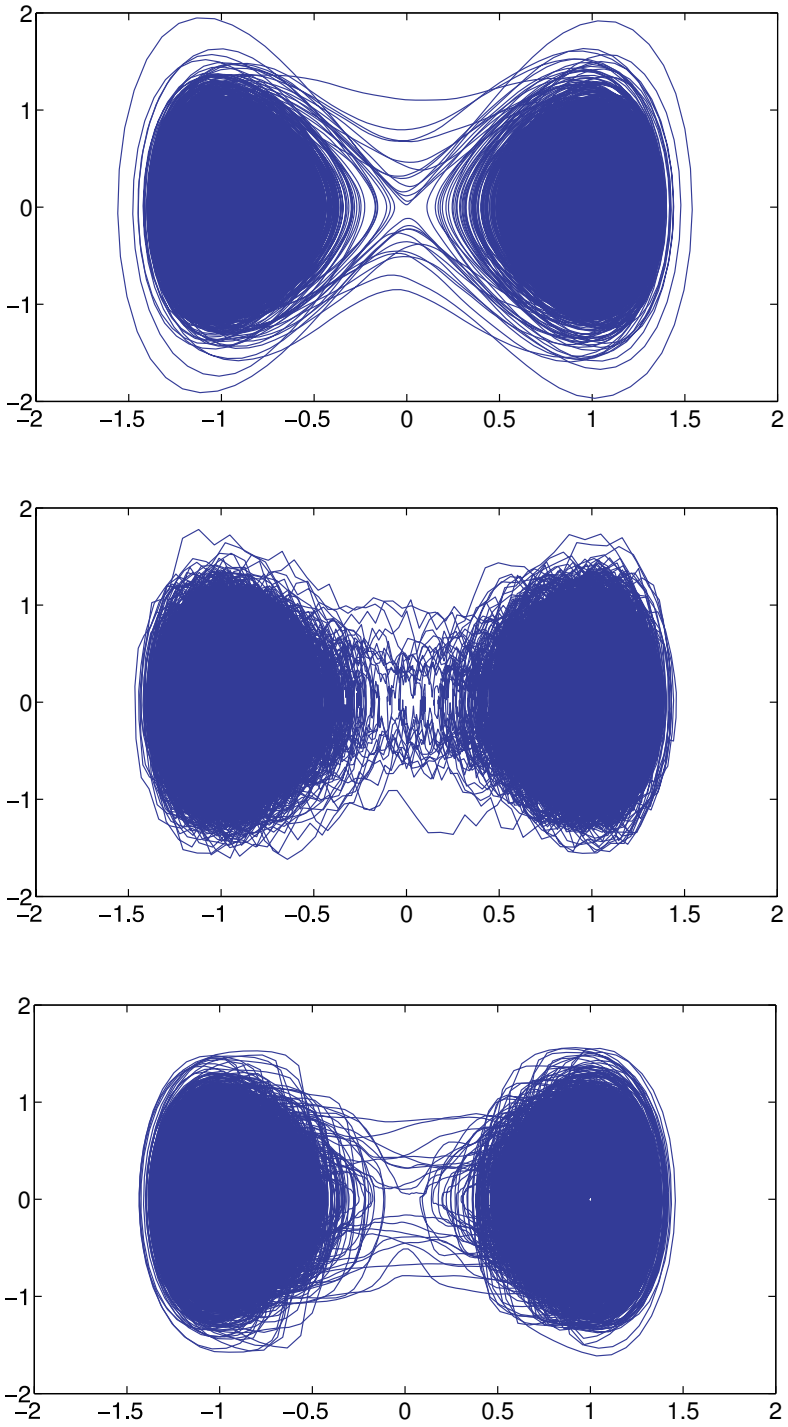

FIG. 2: Trajectories produced by three thermostatting methods for the double well model: Nosé-Hoover-Langevin(top), Langevin dynamics (center), and Embedded Force Method (lower)

.

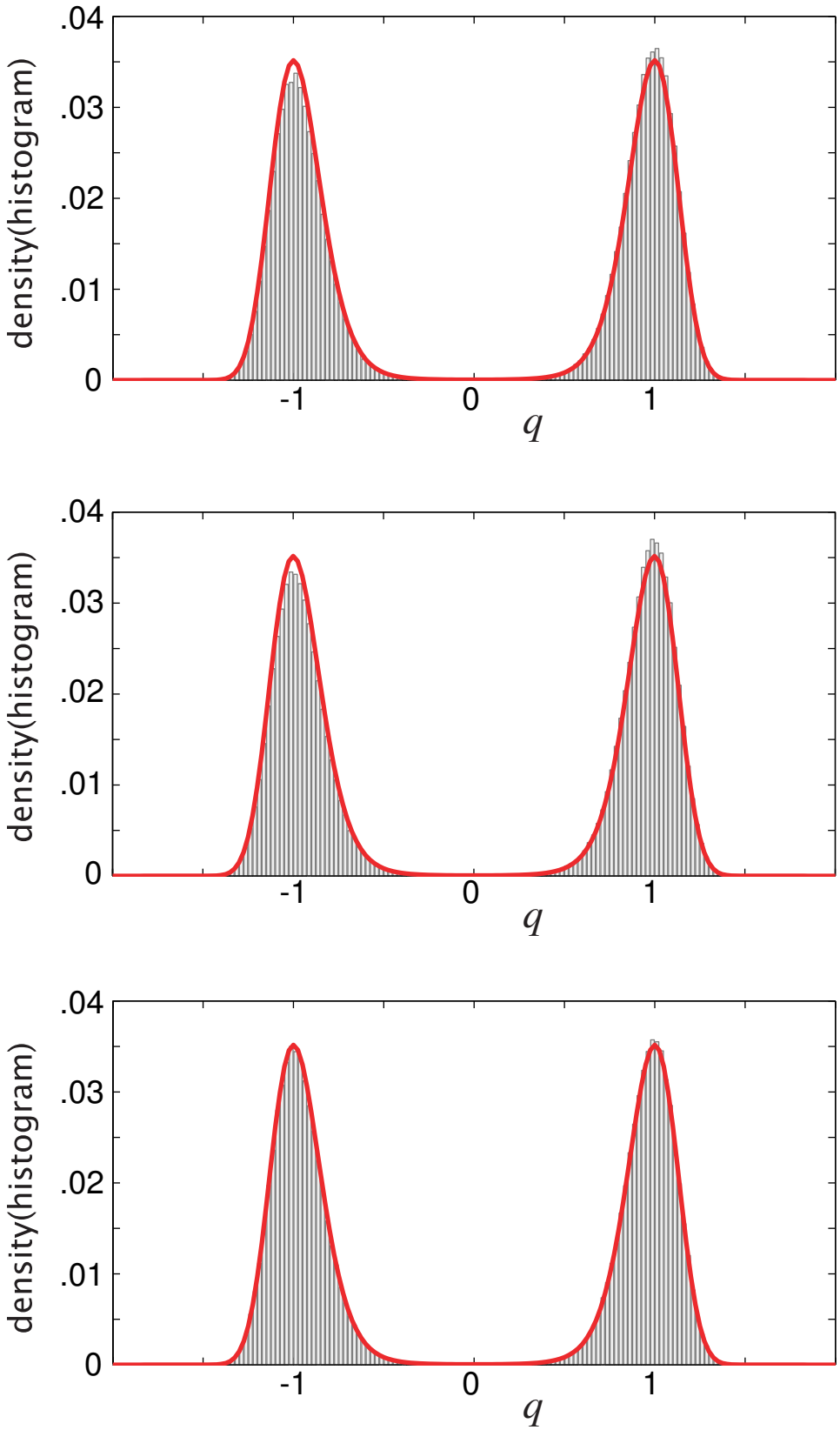

FIG. 3: Sampling of the configurational density using each of the three methods. (NHL: top, Langevin: center, EFM: lower) (Each method exhibits random error, even in 50M steps, so precise comparisons based on these sample trajectories should not be inferred.)

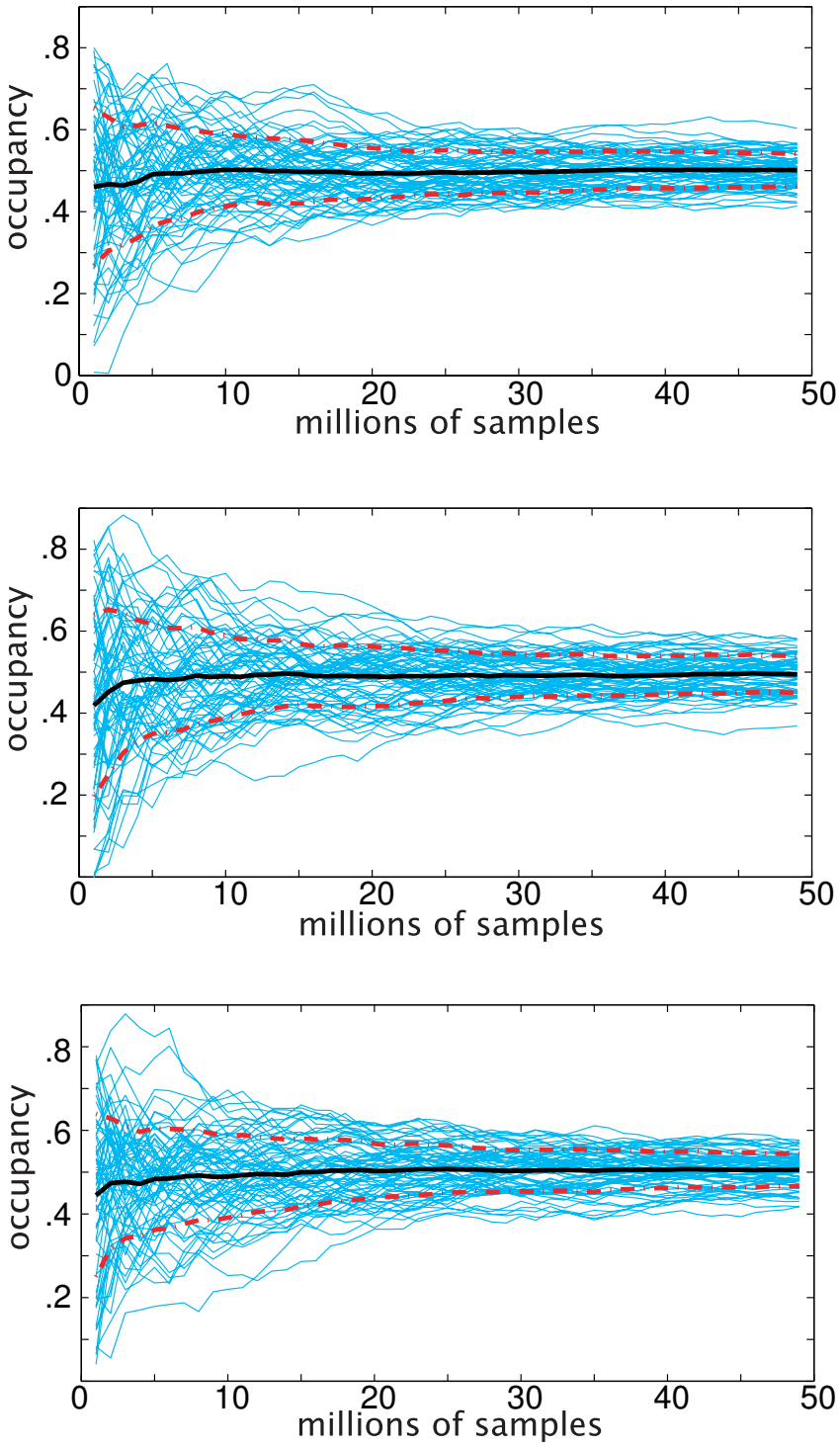

FIG. 4: The occupancy of the right basin should be 1/2. Here we calculate the average occupancy time as a function of sampling steps. (NHL: top, Langevin: center, EFM: lower) The mean is shown as a bold solid line, while the standard error bars are marked by dashed lines.